Information Recovery from Pairwise Measurements

Yuxin Chen
Electrical Engineering, Stanford University
Email: yxchen@stanford.edu

Andrea J. Goldsmith
Electrical Engineering, Stanford University
Email: andrea@wsl.stanford.edu

***Abstract*—A variety of information processing tasks in practice involve recovering n objects from single-shot graph-based measurements, particularly those taken over the edges of some measurement graph $\mathcal{G}$. This paper concerns the situation where each object takes value over a group of M different values, and where one is interested to recover all these values based on observations of certain pairwise relations over $\mathcal{G}$. The imperfection of measurements presents two major challenges for information recovery: 1) *inaccuracy*: a (dominant) portion $1-p$ of measurements are corrupted; 2) *incompleteness*: a significant fraction of pairs are unobservable, i.e. $\mathcal{G}$ can be highly sparse.**

Under a natural random outlier model, we characterize the *minimax recovery rate*, that is, the critical threshold of non-corruption rate p below which exact information recovery is infeasible. This accommodates a very general class of pairwise relations. For various homogeneous random graph models (e.g. Erdős–Rényi random graphs, random geometric graphs, small world graphs), the minimax recovery rate depends almost exclusively on the edge sparsity of the measurement graph $\mathcal{G}$ irrespective of other graphical metrics. This fundamental limit decays with the group size M at a square root rate before entering a connectivity-limited regime. Under the Erdős–Rényi random graph, a tractable combinatorial algorithm is proposed to approach the limit for large M ($M = n^{\Omega(1)}$), while order-optimal recovery is enabled by semidefinite programs in the small M regime.

I. Introduction

Many information processing tasks of interest are concerned with jointly recovering multiple objects from a set of graph-based observations. In various practical scenarios, direct measurement of an isolated object is inaccessible, but one is allowed to measure certain mutual relations across a few object pairs. Such pairwise observations often carry a large amount of information over all objects. As a result, simultaneous recovery of multiple objects becomes feasible as long as the global information consistency across measurements can be appropriately exploited. A small sample of popular applications are described as follows.

Graph-Based Channel Coding. This concerns communication over a set of independent *single-shot* binary channels, where each bit of the codeword is generated as the binary sum of two input information bits associated with an edge [1]. The output of each binary channel forms a noisy observation over an edge. These channels are termed "*graphical channels*", for which the information stability has been investigated by Abbe and Montanari [1], [2].

Joint Graph Matching. Consider n isomophic vertex sets $\mathcal{S}_i$ ($1 \leq i \leq n$). This problem aims to find the global matching of vertices across all these sets. The state-of-the-art procedure [3]–[5] for joint matching is: (1) compute a collection of pairwise maps in isolation; (2) perform global matching based on these noisy pairwise estimates. If one encodes each set by a permutation matrix $\mathbf{\Pi}_i$, then the pairwise matching between $\mathcal{S}_i$ and $\mathcal{S}_j$ can be represented as $\mathbf{\Pi}_j^{-1}\mathbf{\Pi}_i$. This problem arises in numerous applications in computer vision and graphics, DNA / RNA sequencing, solving jigsaw puzzles, etc.

Rotation Registration in Multi-view Reconstruction. Consider n views of a single scene from different angles, represented by n rotation matrices $\boldsymbol{R}_i$ ($1 \leq i \leq n$). One can estimate a collection of relative rotations $\boldsymbol{R}_i^{-1}\boldsymbol{R}_j$ across several pairs of views. The problem of rotation registration aims to recover all rotation matrices $\boldsymbol{R}_i$ (except for some global rigid transform) from the above pairwise estimates. This arises in many applications, e.g. structure from motion [6], finding 3D molecule structure [7], and is also very related to multi-signal alignment and synchronization problems [8].

Robust PCA for Rank-1 Covariance Matrices. Consider a special matrix completion and robust PCA problem [9]–[11] that concerns a rank-1 covariance matrix $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^{\top}$, where $\boldsymbol{x} := [x_1, \cdots, x_n]^{\top}$. Specifically, we obtain estimates of a small sampling of $\boldsymbol{X}_{ij}$, where a dominant portion of samples are maliciously corrupted. The goal is to reconstruct the whole matrix $\boldsymbol{X}$, which is a fundamental tasks in statistical inference.

In practice, pairwise estimates based on an isolated object pair are often imperfect, which presents two major challenges facing information recovery: 1) *inaccuracy*: a significant fraction of obtained measurements are corrupted; 2) *incompleteness*: the measurement graph might be highly sparse. In this paper, we aim at determining the fundamental performance limit of graph-based joint information recovery in the simultaneous presence of these measurement imperfections.

A. Motivation: Lack of Benchmark

While these practical applications have witnessed a flurry of activity in algorithm design, they are based primarily on computational considerations. Inspired by recent success in matrix completion, efficient algorithms have been proposed that can provably tolerate a large outlier rate. If we denote by p the probability that a measurement is not corrupted (termed *non-corruption probability* as defined later in (2)), then the state-of-the-art performance guarantees are summarized below. For simplicity, these results are shown for a complete measurement graph, i.e. measurements across all pairs are available.

- **Joint Graph Matching:** An SDP formulation proposed in [3] provably works in the regime where $p \geq 50\%$, and has been improved to $p \geq \log^2 n/\sqrt{n}$ in [4].
- **Rotation Registration:** For continuous-valued rotation matrices, semidefinite programming (SDP) allows perfect rotation registration as long as $p \geq 50\%$ [12].
- **Rank-1 Robust PCA:** Nuclear norm minimization admits exact disentanglement of a rank-1 matrix from gross outliers as soon as $p = \Omega(\operatorname{poly}\log(n)/\sqrt{n})$ [11].

These algorithms have enjoyed both theoretical and practical success. However, several fundamental questions from an information theoretic view are left unanswered. What is the minimum p that admits accurate joint object recovery? How is this fundamental limit affected by the measurement graph? How does the limit depend on the object representation? Are these fundamental limits achievable by computationally tractable algorithms? We believe that an information theoretic view on the fundamental limits of these algorithms will provide a benchmark for algorithm evaluations and comparisons for practical applications.

Notably, a recent independent work [1] characterized the exact information-theoretic boundary for binary graphical channels associated with the Erdős–Rényi random graphs. However, the fundamental limits under more general graphs remain unknown. Also, our work is inspired by a recent line of work characterizing the fundamental limits of structure detection from matrix-valued measurements [13], [14], although the models therein are quite different from what we study here.

B. Contributions

In this paper, we assume that each object can be represented over a group containing M elements, and consider a very general class of pairwise relations defined over the group. Our main contributions are summarized as follows.

1) Under mild symmetry and cut-set balance assumptions, we develop a fundamental lower limit on the *minimax recovery rate*, that is, the smallest possible non-corruption probability p of measurements below which perfect information recovery is infeasible. This lower limit relies almost exclusively on the maximum vertex degree $d_{\max}$ (or edge sparsity) of the measurement graph $\mathcal{G}$ irrespective of other graphical metrics. The limit exhibits contrasting features in two regimes: 1) when $M = o\left(d_{\max}/\text{poly}\log\left(n\right)\right)$ (called the *information-limited regime*), the bound decays at a rate $\Theta(\sqrt{\text{poly}\log\left(n\right)/(M d_{\max})})$; 2) when $M = \Omega\left(d_{\max}/\text{poly}\log(n)\right)$ (termed the *connectivity-limited regime*), the limit is proportional to $\Theta\left(\text{poly}\log\left(n\right)/d_{\max}\right)$ independent of M.

2) We demonstrate that the above lower limit is tight for a broad class of homogeneous random graphs. Specifically, a maximum compatibility test achieves the derived minimax recovery rate for various graphs, e.g. Erdős–Rényi graphs, random geometric graphs, and small world graphs. For all these graph models, perfect outlier removal is possible even when only a vanishingly small portion of measurements are reliable.

3) Under pairwise difference models, we investigate computational tractability in two special regimes when $\mathcal{G}$ is an Erdős–Rényi graph. For very large M ($M = n^{\Omega(1)}$), we propose a polynomial-time combinatorial algorithm to approach the statistical limit by exploiting global cycle consistency. In the small M regime ($M = O(1)$), our findings reveal that order-optimal information recovery is achievable by SDP.

C. Terminology and Notations

For ease of presentation, we introduce some graph theory terminology as follows. For any vertex set S, denote by ∂S the *edge boundary* of S, which is the set of all edges with exactly one endpoint in S. For any vertex sets $S_1, S_2 \subseteq V$, we let $\mathcal{E}_{S_1,S_2}$ denote the set of edges crossing from S_1 to S_2. A cycle $\mathcal{C}$ is called a k-cycle if it contains k edges.

Several random graph models are defined below; see [15] for an introduction. An Erdős–Rényi graph of n vertices, denoted by $\mathcal{G}_{n,q}$, is constructed in a way such that each pair of vertices is independently connected by an edge with probability q. A random geometric graph, denoted by $\mathcal{G}_{n,r}$, is generated via a 2-step procedure: i) place n vertices at random uniformly and independently on the surface of a unit sphere; ii) connect two vertices by an edge if the Euclidean distance between them is at most r. An expander graph with edge expansion $h_{\mathcal{G}}$ satisfies $|\partial S| \geq h_{\mathcal{G}}|S|$ for all vertex sets S satisfying $|S| \leq n/2$. A (Watts and Strogatz) small-world graph with degree k and rewiring probability q is constructed as follows: start from a ring with n vertices and k edges per vertex, then rewire each edge with probability q by connecting one end to a randomly chosen vertex.

II. Problem Formulation

In this section, we describe the object representation and measurement models in a precise manner.

Object Representation. Consider a large network of n vertices $\mathcal{V} = \{1, \cdots, n\}$, each taking a value x_i. Here, x_i's can be finite-bit strings, vectors, or any others that can be represented over a *group*. Formally, suppose that $x_i \in \mathbb{G}_M$, where $\mathbb{G}_M$ represents a group (together with an operation "+") that contains $M \geq 2$ elements (denoted by $\{0, 1, \cdots, M-1\}$). We will denote by $-x$ the inverse element of x.

Pairwise Relation. Denote by $x_i \ominus x_j$ the pairwise relation of interest between x_i and x_j, which is a binary operation $\mathbb{G}_M \times \mathbb{G}_M \mapsto \mathbb{G}_M$. We assume throughout that "$\ominus$" satisfies the following property: for any $x_1, x_2 \in \mathbb{G}_M$,

$$\begin{cases} x_1 \ominus x_2 \neq x_1 \ominus x_3, & \forall x_2 \neq x_3; \\ x_1 \ominus x_2 \neq x_4 \ominus x_2, & \forall x_1 \neq x_4. \end{cases} \tag{1}$$

Intuitively, this ensures that "$\ominus$" yields a sufficient amount of difference whenever any side of "$\ominus$" is perturbed, which in turn allows us to detect the perturbation. For notational convenience, we write

$$\boldsymbol{x} \ominus \boldsymbol{x} := \left[x_i \ominus x_j\right]_{1 \leq i,j \leq n}. \tag{2}$$

The above definition subsumes a very broad class of pairwise relations. For example, in the conventional arithmetic finite field, any function $x_i \ominus x_j := a x_i + b x_j \pmod{M}$ falls within this class as long as both (a, M) and (b, M) are coprime. The most commonly encountered relations are the pairwise difference $x_i \ominus x_j := x_i + (-x_j)$ and pairwise sum $x_i \ominus x_j := x_i + x_j$. Some practical examples are as follows:

1) **Pairwise matching** $\boldsymbol{\Pi}_i^{-1}\boldsymbol{\Pi}_j$: if we set $\mathbb{G}_M = \{\text{all permutation matrices } \boldsymbol{\Pi}\}$ and let "+" denote matrix multiplication, then $\boldsymbol{\Pi}_i^{-1}\boldsymbol{\Pi}_j$ is also pairwise difference.
2) **Relative rotation** $\boldsymbol{R}_i^{-1}\boldsymbol{R}_j$: if we set $\mathbb{G}_M = \{\text{all rotation matrices } \boldsymbol{R}\}$ and let "+" denote matrix multiplication, then $\boldsymbol{R}_i^{-1}\boldsymbol{R}_j$ can be treated as a special case of pairwise difference.
3) **Entry multiplication** $x_i x_j$: if we set $\mathbb{G}_M = \{\log x : \ x \in \mathbb{S}\}$ for some set $\mathbb{S}$ and let "+" denote

arithmetic addition, then $\log(x_i x_j) = \log x_i + \log x_j$ is subsumed by the pairwise sum model.

Measurement Model. A measurement graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ comprises $\mathcal{V}$ together with a set $\mathcal{E}$ of undirected edges, with $d_{\min}$ and $d_{\max}$ denoting the minimum and maximum vertex degrees, respectively. We are allowed to obtain measurements $\boldsymbol{y} = [y_{ij}]_{(i,j)\in\mathcal{G}}$ of the relation between x_i and x_j over all $(i,j) \in \mathcal{G}$. Each measurement is independently corrupted following a distribution

$$\mathbb{P}_{y_{ij}} = p\delta_{x_i \ominus x_j} + (1-p)\mathrm{Unif}_M, \tag{3}$$

where δ_x denotes the Dirac measure on the point x, and Unif_M represents the uniform distribution over all M elements. In other words, a fraction $(1-p)$ of measurements behave as *random outliers* and are unreliable. In contrast to conventional information theory settings, all measurements are *single-shot* with no channel reuse.

We say that *perfect information recovery* is achievable by a method ψ if, with vanishing probability of error when $n \to \infty$, ψ recovers *all* x_i's (except for some unrecoverable global state) or, equivalently, the entire matrix $\boldsymbol{x} \ominus \boldsymbol{x}$. We are interested in identifying the *rate* of the following quantity:

- **Minimax Recovery Rate** $p^*(\mathcal{G}, M)$: the smallest non-corruption rate p below which no method allows perfect information recovery simultaneously for all inputs.

Our goal is to determine the fundamental tradeoff among $p^*(\mathcal{G}, M)$, the group size M, and the properties of $\mathcal{G}$.

III. Main Results

This section presents the main results of this paper, characterizing $p^*(\mathcal{G}, M)$ for a large class of homogeneous graphs, including the most widely studied models like Erdős–Rényi graphs, random geometric graphs, small world graphs, etc.

A. Preliminaries: Key Graphical Metrics

Before proceeding to our main results, we introduce a couple of graphical quantities of $\mathcal{G}$. For any integer k, define

$$\mathcal{N}_k := \{S \subseteq V : |\partial S| \leq k\}, \text{ and } N_k = |\mathcal{N}_k|, \tag{4}$$

whereas the exponents of N_k are define as follows

$$\begin{aligned} \forall k:\ \alpha_k^{\mathrm{lb}} &:= \log\left(N_{k d_{\min}}\right)/k, \qquad \alpha_k^{\mathrm{ub}} := \log\left(N_{k d_{\max}}\right)/k, \\ \alpha_{\mathrm{m}}^{\mathrm{lb}} &:= \max\nolimits_k \alpha_k^{\mathrm{lb}}, \qquad \alpha_{\mathrm{m}}^{\mathrm{ub}} := \max\nolimits_k \alpha_k^{\mathrm{ub}}. \end{aligned} \tag{5}$$

Intuitively, for a homogeneous graph (take a complete graph as an extreme case), the number of cut-sets containing at most $kd_{\max}$ edges is expected to grow exponentially with k when $k = O(n^2/d_{\max})$. In light of this, $\alpha_{\mathrm{m}}^{\mathrm{lb}}$ and $\alpha_{\mathrm{m}}^{\mathrm{ub}}$ quantify *the degree of unbalance* in terms of the cut-set distribution.

Besides, for some large enough constant K, we define

$$\beta_{\mathrm{m}}^K := \sup_{S \subsetneq V:\ \frac{|\partial S|}{|S| d_{\min}} \leq K} \left| \left\{ S_1 \subseteq S:\ \frac{\left|\mathcal{E}_{S_1 \times S \backslash S_1}\right|}{|S|\, d_{\min}} \geq \frac{K-3}{K} \right\} \right|.$$

In other words, β_{m}^K evaluates the possibility of two disjoint subsets S_1, S_2 such that a dominant portion of edges incident to the vertex set $S_1 \cup S_2$ reside within the edge set $\mathcal{E}_{S_1,S_2}$.

Interestingly, for various homogeneous random graphs, $\alpha_{\mathrm{m}}^{\mathrm{lb}}$, $\alpha_{\mathrm{m}}^{\mathrm{ub}}$, and β_{m}^K are all very small, as stated below.

Lemma 1. *(1)* ***Expander graphs****. If $\mathcal{G}$ is an expander graph with edge expansion $h_{\mathcal{G}}$ (defined in Section I-C), then $\alpha_{\mathrm{m}}^{\mathrm{lb}}, \alpha_{\mathrm{m}}^{\mathrm{ub}} = O\left(\frac{d_{\max}}{h_{\mathcal{G}}}\log n\right)$. Besides, $\beta_{\mathrm{m}}^K = 0$ for any $K > \frac{d_{\max}}{h_{\mathcal{G}}}$.*

(2) ***Geometric graphs****. Suppose that there exists a universal constant $c_0 > 0$ such that for any vertex pair $(i,j) \in \mathcal{G}$, $|\partial\{i\} \cap \partial\{j\}| \geq c_0 d_{\min}$. Then $\alpha_{\mathrm{m}}^{\mathrm{lb}}, \alpha_{\mathrm{m}}^{\mathrm{ub}} = O(\log^2 n)$. Also, there is a large constant $K > 0$ such that $\beta_{\mathrm{m}}^K = 0$.*

B. Minimax Recovery Rate

Encouragingly, for a broad class of homogeneous graph models, the fundamental minimax rate $p^*(\mathcal{G}, M)$ depends almost only on the group size M and the edge sparsity of $\mathcal{G}$ irrespectively of any other graphical properties, as stated in the following theorem.

Theorem 1. *Suppose that the min-cut size of graph $\mathcal{G}$ is at least $c_0 d_{\min}$ for some constant $c_0 > 0$ and that $\mathcal{G}$ satisfies*

$$d_{\max}/d_{\min} = \Theta\left(\mathrm{poly}\log(n)\right), \tag{6}$$

$$\alpha_{\mathrm{m}}^{\mathrm{lb}}, \alpha_{\mathrm{m}}^{\mathrm{ub}} = O\left(\mathrm{poly}\log\left(n\right)\right), \tag{7}$$

$$\beta_{\mathrm{m}}^K = n^{O(\mathrm{poly}\log(n))}. \tag{8}$$

for some $K = \mathrm{poly}\log(n)$. Then one has

$$p^*(\mathcal{G}, M) = \begin{cases} \Theta\left(\sqrt{\frac{\mathrm{poly}\log(n)}{d_{\max} M}}\right), & \text{if } M = O\left(\frac{d_{\max}}{\mathrm{poly}\log(n)}\right), \\ \Theta\left(\frac{\mathrm{poly}\log(n)}{d_{\max}}\right), & \text{if } M = \Omega\left(\frac{d_{\max}}{\mathrm{poly}\log(n)}\right). \end{cases} \tag{9}$$

Theorem 1 characterizes the minimax recovery rate – the lowest possible non-corruption rate that admits perfect information recovery, provided that the measurement graph $\mathcal{G}$ satisfies certain assumptions on degree homogeneity and cut-set balance. The minimax rate (9) relies primarily on the following two features.

- **Edge Sparsity**: When (6) is satisfied, for each vertex, a total number of $\Theta\left(d_{\max}\right)$ edges are observable out of $n-1$ possible pairs, which characterizes the undersampling ratio of the entire matrix-valued signal. Somewhat unexpectedly, the minimax limit (9) relies almost exclusively on the edge sparsity (or $d_{\max}$) irrespectively of any other graphical metrics.
- **Information Carried by Each Object**: For a group of cardinality M, the information contained in each object is determined by M. Interestingly, the minimax recovery rate (9) decays with M before hitting a fundamental connectivity barrier $\Theta(\mathrm{poly}\log\left(n\right)/d_{\max})$.

The fundamental limits (9) for the two regimes exhibit contrasting features, which we discuss separately as follows.

i. Information-limited regime ($M = o\left(\frac{d_{\max}}{\mathrm{poly}\log(n)}\right)$): when M is small, the amount of information that can be conveyed through each pairwise measurement is limited. This results in many false negatives, namely, a large number of incorrect patterns $\hat{\boldsymbol{x}} \neq \boldsymbol{x}$ are not statistically distinguishable from $\boldsymbol{x}$ when p is very small. In this regime, the fundamental limit (9) decays with M and $d_{\max}$ both at square-root rates.

ii. Connectivity-limited regime ($M = \Omega\left(\frac{d_{\max}}{\mathrm{poly}\log(n)}\right)$): as M grows, we obtain sufficient information to preclude

	$M=O\left(\frac{np_{\mathrm{obs}}}{\log n}\right)$	$\omega\left(\frac{np_{\mathrm{obs}}}{\log n}\right)\leq M\leq\Theta\left(np_{\mathrm{obs}}\right)$	$M=\omega\left(np_{\mathrm{obs}}\right)$
$p_{\min}$	$\Theta\left(\sqrt{\frac{\log n}{np_{\mathrm{obs}}M}}\right)$	$\Theta\left(\frac{\log n}{np_{\mathrm{obs}}\log\left(\frac{2M\log n}{d_{\max}}\right)}\right)$	$\Theta\left(\frac{\log n}{np_{\mathrm{obs}}}\right)$

(a) Erdős–Rényi graphs $\mathcal{G}_{n,p_{\mathrm{obs}}}$ ($p_{\mathrm{obs}}=\omega\left(\frac{\log n}{n}\right)$)

	$M=O\left(\frac{nr^2}{\log n}\right)$	$\omega\left(\frac{nr^2}{\log n}\right)\leq M\leq\Theta\left(nr^2\right)$	$M=\omega\left(nr^2\right)$
$p_{\min}$	$\Theta\left(\sqrt{\frac{\mathrm{poly}\log(n)}{nr^2M}}\right)$	$\Theta\left(\frac{\mathrm{poly}\log(n)}{nr^2\log\left(\frac{2M\log n}{nr^2}\right)}\right)$	$\Theta\left(\frac{\log n}{nr^2}\right)$

(b) Random geometric graphs $\mathcal{G}_{n,r}$ ($r=\omega\left(\sqrt{\frac{\log n}{n}}\right)$)

Table I
MINIMAX RECOVERY RATES FOR TWO RANDOM GRAPHS.

more false negatives. However, the measurement graph $\mathcal{G}$ presents another fundamental connectivity bottleneck. In fact, if $p=O(1/d_{\max})$, then there exists at least one vertex that is not linked with any reliable measurement, and hence no information on this isolated object can be conveyed through the obtained pairwise measurements.

When the min-cut size of $\mathcal{G}$ is large (i.e. $\omega\left(\log n\right)$), Theorem 1 implies that accurate information recovery is possible even when a dominant portion of measurements are outliers.

Theorem 1 is applicable to a very broad class of measurement patterns, which subsumes many random graph models. Some of the most widely used ones are summarized below.

1) Erdős–Rényi graphs. If $\mathcal{G}\sim\mathcal{G}_{n,p_{\mathrm{obs}}}$ for some observation probability $p_{\mathrm{obs}}>\frac{\log n}{n}$, then $p^{*}(\mathcal{G},M)$ decays with $\Theta\left(\sqrt{\frac{\log(n)}{np_{\mathrm{obs}}M}}\right)$ before it hits the barrier $\Theta\left(\frac{\log n}{np_{\mathrm{obs}}}\right)$.

2) Random Geometric Graphs. If $\mathcal{G}\sim\mathcal{G}_{n,r}$ for some ratio $r=\Omega\left(\frac{\log n}{\sqrt{n}}\right)$, then $p^{*}(\mathcal{G},M)$ vanishes with $\Theta\left(\frac{\mathrm{poly}\log(n)}{\sqrt{nr^2M}}\right)$ before reaching the connectivity threshold $\Theta\left(\frac{\log n}{nr^2}\right)$.

3) Small-World Graphs. If $\mathcal{G}$ is constructed according to the Watts and Strogatz model with degree k and some constant rewiring probability $q>0$, then $p^{*}(\mathcal{G},M)$ vanishes with $\Theta\left(\frac{\mathrm{poly}\log(n)}{\sqrt{kM}}\right)$ before entering the connectivity-limited regime $\Theta\left(\frac{\mathrm{poly}\log(n)}{k}\right)$.

More refined characterizations of Erdős–Rényi graphs and random geometric graphs are summarized in Table I.

In the following subsections, we develop a lower bound on p via the generalized Fano's inequality [16], as well as a recovering strategy that achieves the minimax recovery rates. These results taken collectively establish Theorem 2.

C. Fundamental Lower Bound

Our converse results are divided into 3 regimes, as formally stated in the following theorem.

Theorem 2. *For any small constant $\epsilon>0$, there does not exists any method that allows asymptotically perfect information recovery unless*

$$p>\begin{cases}\sqrt{\frac{(1-\epsilon)\log((M-1)n)}{(d_{\max}+1)(M-1)}}, & \text{if } M=o\left(\frac{d_{\max}}{\log(Mn)}\right),\\ \frac{(1-\epsilon)\log((M-1)n)}{d_{\max}\log\left(\frac{2M\log(Mn)}{d_{\max}}\right)}, & \text{if } \Theta\left(\frac{d_{\max}}{\log(Mn)}\right)\leq M\leq\Theta\left(d_{\max}\right)\\ \frac{1}{d_{\max}}, & \text{if } M=\omega\left(d_{\max}\right).\end{cases}\tag{10}$$

Note that there exists a narrow transition regime ($\Theta(d_{\max}/\alpha_{\mathrm{m}}^{\mathrm{ub}})\leq M\leq\Theta(d_{\max})$) where the decaying rate with M has rapidly transited from $\Theta(1/\sqrt{M})$ to $\Theta(1)$ within this small interval.

The proof is based on the generalized Fano's inequality [16], and is deferred to [17]. Somewhat unexpectedly, conditional on the ground truth $\boldsymbol{x}=\boldsymbol{0}$, most alternative hypotheses $\boldsymbol{x}\neq\boldsymbol{0}$ are equally difficult to distinguish. For a large class of graphs, this feature arises independent of precise graphical properties.

D. Optimal Recovery Procedure

We are now ready to introduce the recovery method ψ, which is a maximum compatibility test defined by

$$\psi\left(\boldsymbol{y}\right):=\arg\max_{\boldsymbol{x}\in\{0,1,\cdots,M-1\}^n}\sum_{(i,j)\in\mathbb{G}}\mathbb{I}\left(y_{i,j}=\left[\boldsymbol{x}\ominus\boldsymbol{x}\right]_{i,j}\right),\tag{11}$$

where $\boldsymbol{y}=\left[y_{ij}\right]_{1\leq i,j\leq n}$ represents the observation. Encouragingly, for a large class of measurement graphs, this recovery rule admits perfect information recovery as soon as the non-corruption rate exceeds the lower bound (10), as revealed in the following theorem.

Theorem 3. *Suppose that the min-cut of $\mathcal{G}$ contains at least $c_0d_{\min}$ cut edges*[1] *for some absolute constant $c_0>0$. There exist universal constant $c_1,c_2,K>0$ such that if either of the following is satisfied:*

i) $M=o\left(\frac{d_{\min}}{\log n+\alpha_{\mathrm{m}}^{\mathrm{lb}}}\right)$, and

$$p\geq c_1\sqrt{\frac{\log n+\alpha_{\mathrm{m}}^{\mathrm{lb}}+\log\beta_{\mathrm{m}}^{K}}{Md_{\min}}};\tag{12}$$

ii) $M=\omega\left(d_{\min}\right)$ and

$$p>\frac{c_1\left(\log n+\alpha_{\mathrm{m}}^{\mathrm{lb}}+\log\beta_{\mathrm{m}}^{K}\right)}{d_{\min}};\tag{13}$$

iii) $M=c\log c\cdot\frac{d_{\min}}{\log n+\alpha_{\mathrm{m}}^{\mathrm{lb}}}$ for some $c\log c=O\left(\log n\right)$ and $c\log c=\Omega\left(1\right)$, and

$$p>\frac{c_1\max\left\{\log n,\sqrt{\alpha_{\mathrm{m}}^{\mathrm{lb}}\log n}\right\}}{d_{\min}\sqrt{\log\left(\frac{M\left(\log n+\alpha_{\mathrm{m}}^{\mathrm{lb}}+\log\beta_{\mathrm{m}}^{K}\right)}{d_{\min}}\right)}};\tag{14}$$

then $\psi\left(\cdot\right)$ defined in (11) admits exact information recovery for all inputs with probability exceeding $1-c_2n^{-3}$.

Theorem 3 quantifies the regimes of p such that $\psi(\cdot)$ admits reliable information recovery. The boundaries of these regimes coincide with the lower bound derived in Theorem 2 as long as $d_{\min}=\Theta(d_{\max})$, $\beta_{m}^{K}=n^{O(1)}$, which can be satisfied for a very large class of homogeneous graphs.

[1]For a broad class of homogeneous random graphs (e.g. Erdős–Rényi graph, random geometric graphs, rings), this assumption can be easily satisfied.

IV. Tractable Algorithms for Two Regimes

Now that we have characterized the fundamental minimax limits on fault tolerance rate, a natural question arises as to whether there is a price to pay for computational tractability. In general, to answer this question, algorithms need to be studied on a case-by-case basis. For concreteness, we focus on Erdős–Rényi graphs, i.e. $\mathcal{G} \sim \mathcal{G}_{n,p_{\mathrm{obs}}}$, and the pairwise differential model $x_i \ominus x_j = x_i + (-x_j)$ under certain group operation "+". These are arguably among the most important models in practice.

A. Large M Regime: $M = n^{\Omega(1)}$

Interestingly, when M is sufficiently large, there exist tractable algorithms that can approach the minimax recovery rates. Note that the measurement pattern imposes many global compatibility requirements, and the key is to exploit these consistency constraints.

For the pairwise differential model, the most important global constraint can be summarized as "cycle-consistency", which states that for all cycles $\mathcal{C}$ in the graph $\mathcal{G}$, the sum of pairwise difference over all edges along the cycle is 0, i.e.

$$\sum\nolimits_{(i,j)\in\mathcal{C}} x_i \ominus x_j = 0. \tag{15}$$

This criterion can be exploited to detect and prune outliers. As a result, we propose a tractable[2] combinatorial algorithm by searching over several classes of cycles, as summarized in Algorithm 1. It turns out that this polynomial-time algorithm approaches the minimax limit as $\log M$ increases.

Algorithm 1 Polynomial-time algorithm when $M = n^{\Omega(1)}$.

Input: measurements over $\mathcal{G}$, and a given cycle order k.

1) Find all edges contained in at least 1 *zero-sum* k-cycle in $\mathcal{G}$. Denote the set of these edges by $\mathcal{E}_{\mathrm{true}}^{k}$.

2) For any $i \neq j$, pick a path $\mathcal{P}_{i,j}$ in $\mathcal{E}_{\mathrm{true}}^{k}$ connecting i and j. Set $x_i \ominus x_j = \sum_{(k,l)\in\mathcal{P}_{i,j}} x_k \ominus x_l$.

The following theorem characterizes the regime where Algorithm 1 admits perfect information recovery.

Theorem 4. *Suppose that the measurement graph $\mathcal{G}$ is drawn from an Erdős–Rényi model $\mathcal{G}_{n,p_{\mathrm{obs}}}$ for some quantity p_{obs}. For any given cycle order k $(3 \leq k \leq \log n)$ and any small constant $\epsilon > 0$, Algorithm 1 admits perfect recovery with probability exceeding $1 - \frac{1}{n^{\epsilon}}$, provided that $M = n^{k+\epsilon}$ and*

$$p \geq \frac{c_1 \log^2 n}{p_{\mathrm{obs}} n^{\frac{k-2}{k-1}}} \tag{16}$$

for some universal constant $c_1 > 0$.

Theorem 4 indicates that, for any small $\delta > 0$, Algorithm 1 admits perfect information recovery as long as

$$p = \Omega\left(\log^2 n / \left(p_{\mathrm{obs}} n^{1-\delta}\right)\right), \quad \text{and} \quad M \geq n^{\frac{1}{\delta}+1+o(1)}.$$

Consequently, for very large M, there exist tractable algorithms that can approach the fundamental limits arbitrarily closely.

[2] Note that M can often be encoded by $\log M$ bits, and usually an elementary operation takes polynomial time as long as $\log M = \mathrm{poly}(n)$.

B. Constant M Regime: $M = O(1)$

In the regime where $M = O(1)$, the fundamental limit reverts to $p_{\min} = \Theta\left(\sqrt{\frac{\log n}{n p_{\mathrm{obs}}}}\right)$. While the algorithms need to be designed based on the precise definition of the operation "+", in may situations appropriate SDP formulations are able to achieve the limit. For instance, in the joint object matching or rotation registration case, perfect information recovery is achievable by SDP as soon as

$$p = \Omega\left(\frac{\log^2(n)}{\sqrt{n p_{\mathrm{obs}}}}\right).$$

as detailed in [4]. In fact, SDP is also optimal for other applications like rotation registration. This phenomenon has also been independently observed by Abbe et. al. [1] for the binary-valued case ($M = 2$) with precise preconstants.

V. Conclusion

This paper investigates recovery of multiple objects based on noisy graph-based measurements, which has found numerous applications. For various symmetric random graph models, we have identified the critical fault tolerance boundary beyond which no method allows perfect information recovery. We expect that such fundamental limits provide a benchmark for evaluating the performance of practical algorithms. In addition, we propose tractable combinatorial algorithms that approach the minimax rates for large M, and show that SDP achieves order-optimal fault-tolerance rate in the small M regime.

It remains to be seen how the minimax recovery rates are affected by more general noise models and more general pairwise relations (e.g. the ones that occur in general matrix completion problems). It would also be of great interest to investigate computational tractability for the entire regime, i.e. whether there is a computational gap from statistical limits.

Acknowledgment

The authors would like to thank A. Montanari and Q. Huang for stimulating discussions. This work is supported in part by the NSF Center for Science of Information under grant NSF-CCF- 0939370 and by AFOSR Grant FA9550-09-1-0643.

References

[1] E. Abbe and A. Montanari, "The mutual information of a class of graphical channels," in *Allerton*, Oct 2013, pp. 20–25.

[2] ——, "Conditional random fields, planted constraint satisfaction and entropy concentration," in *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, 2013, pp. 332–346.

[3] Q. Huang and L. Guibas, "Consistent shape maps via semidefinite programming," in *SGP*, vol. 32, no. 5, 2013, pp. 177–186.

[4] Y. Chen, L. Guibas, and Q. Huang, "Near-optimal joint object matching via convex relaxation," *International Conference on Machine Learning (ICML)*, June 2014. [Online]. Available: http://arxiv.org/abs/1402.1473

[5] D. Pachauri, R. Kondor, and V. Singh, "Solving the multi-way matching problem by permutation synchronization," *NIPS*, December 2013.

[6] D. Crandall, A. Owens, N. Snavely, and D. Huttenlocher, "SfM with MRFs: discrete-continuous optimization for large-scale structure from motion," *IEEE CVPR*, pp. 3001–3008, 2011.

[7] M. Cucuringu, A. Singer, and D. Cowburn, "Eigenvector synchronization, graph rigidity and the molecule problem," *Information and Inference*, vol. 1, no. 1, pp. 21–67, 2012.

[8] A. Bandeira, M. Charikar, A. Singer, and A. Zhu, "Multireference alignment using semidefinite programming," *ITCS*, pp. 459–470, 2014.

[9] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of ACM*, vol. 58, no. 3, pp. 11:1–11:37, Jun 2011.

[10] A. Ganesh, J. Wright, X. Li, E. J. Candes, and Y. Ma, "Dense error correction for low-rank matrices via principal component pursuit," in *IEEE ISIT*. IEEE, 2010, pp. 1513–1517.

[11] Y. Chen, A. Jalali, S. Sanghavi, and C. Caramanis, "Low-rank matrix recovery from errors and erasures," *IEEE Trans on Info Theory*, vol. 59, no. 7, pp. 4324–4337, 2013.

[12] L. Wang and A. Singer, "Exact and stable recovery of rotations for robust synchronization," *arXiv preprint arXiv:1211.2441*, 2012.

[13] Z. Ma and Y. Wu, "Computational barriers in minimax submatrix detection," *arXiv preprint arXiv:1309.5914*, 2013.

[14] M. Kolar, S. Balakrishnan, A. Rinaldo, and A. Singh, "Minimax localization of structural information in large noisy matrices," in *Advances in Neural Information Processing Systems*, 2011, pp. 909–917.

[15] D. Shah, *Gossip algorithms*. Now Publishers Inc, 2009.

[16] S. Verdu *et al.*, "Generalizing the Fano inequality," *IEEE Transactions on Information Theory*, vol. 40, no. 4, pp. 1247–1251, 1994.

[17] Y. Chen and A. J. Goldsmith, "Joint information recovery from graph-based measurements," *in preparation*, 2014.